A Genetic Tuning to Improve the Performance of Fuzzy Rule-Based Classification Systems with Interval-Valued Fuzzy Sets: Degree of Ignorance and Lateral Position

J. Sanz[a], A. Fernández[b], H. Bustince[a], F. Herrera[c]

[a] *Departamento de Automática y Computación, Universidad Pública de Navarra, Campus Arrosadía s/n, P.O. Box 31006, Pamplona, Spain*
[b] *Department of Computer Science, University of Jaén, P.O. Box 23071, Jaén, Spain*
[c] *Dept. of Computer Science and Artificial Intelligence, CITIC-UGR (Research Center on Information and Communications Technology), University of Granada, P.O. Box 18071, Granada, Spain*

Abstract

Fuzzy Rule-Based Systems are appropriate tools to deal with classification problems due to their good properties. However, they can suffer a lack of system accuracy as a result of the uncertainty inherent in the definition of the membership functions and the limitation of the homogeneous distribution of the linguistic labels.

The aim of the paper is to improve the performance of Fuzzy Rule-Based Classification Systems by means of the Theory of Interval-Valued Fuzzy Sets and a post-processing genetic tuning step. In order to build the Interval-Valued Fuzzy Sets we define a new function called weak ignorance for modeling the uncertainty associated with the definition of the membership functions. Next, we adapt the fuzzy partitions to the problem in an optimal way through a cooperative evolutionary tuning in which we handle both the degree of ignorance and the lateral position (based on the 2-tuples fuzzy linguistic representation) of the linguistic labels.

The experimental study is carried out over a large collection of data-sets and it is supported by a statistical analysis. Our results show empirically that the use of our methodology outperforms the initial Fuzzy-Rule Based Classification System. The application of our cooperative tuning enhances the results provided by the use of the isolated tuning approaches and also improves the behavior of the genetic tuning based on the 3-tuples fuzzy linguistic representation.

Keywords: Fuzzy Rule-Based Classification Systems; Interval-Valued Fuzzy Sets; Ignorance Functions; Linguistic 2-tuples representation; Genetic Fuzzy Systems; Tuning; Genetic Algorithms.

Email addresses: `joseantonio.sanz@unavarra.es` (J. Sanz), `alberto.fernandez@ujaen.es` (A. Fernández), `bustince@unavarra.es` (H. Bustince), `herrera@decsai.ugr.es` (F. Herrera)

1. Introduction

Computational Intelligence based methods have shown themselves to be useful tools to solve complex problems in classification tasks. Among them, Fuzzy Rule-Based Classification Systems (FRBCSs) [36] are widely employed since they allow us to deal with noisy, imprecise or incomplete information which is often present in many real world problems. They provide a good trade-off between the empirical precision of traditional engineering techniques and the interpretability achieved through the use of linguistic labels whose semantic is close to the natural language. The high number of real world applications (see [1, 42, 50, 57]) in which these systems are employed supports their goodness in dealing with classification problems.

Hybrid approaches are commonly applied to improve the behavior of fuzzy systems [8, 20, 24, 46]. One of the most popular approaches is the hybridization of fuzzy logic and Genetic Algorithms (GAs) leading to Genetic Fuzzy Systems (GFSs) [17, 18, 33]. A GFS is basically a fuzzy system augmented by a learning process based on evolutionary computation. The most extended GFS type is the genetic fuzzy rule-based system, where an evolutionary algorithm is employed to learn or tune different components of a fuzzy rule-based system. Specifically, the genetic tuning process consists of automatically selecting the best system parameters for improving the performance of the final model without modifying the existing rule base. The applicability of genetic tuning is clearly shown in the specialized literature, where we can find proposals that introduce linguistic modifiers for tuning the membership functions [12], approaches to perform lateral tuning [5, 25], multiobjective approaches for rule reduction and parameter tuning [26], methods for tuning of Type-2 fuzzy systems [13, 34, 43, 55] or applications to solve real world problems [45], among others.

The success of the methods previously presented depends, to a large degree, on the choice of the membership functions which model the linguistic labels. This choice is a complex problem as there is an inherent uncertainty in the definition of the membership functions, which are usually defined either homogeneously over the input space or by means of expert knowledge. In both cases, the uncertainty can be taken into account by employing the theory of Interval-Valued Fuzzy Sets (IVFSs) [48].

An IVFS provides a lower and an upper bound for the membership value. It defines a membership interval whose length can be seen as the degree of uncertainty when assigning the membership of the element to the set. In fact, one of the most recent construction methods of IVFSs is based on ignorance functions in order to model this uncertainty [11]. IVFSs have been applied successfully in classification tasks [49], approximate reasoning [9], decision making [14] and image processing [10], among others.

Following on from the above, we propose a methodology in which we use IVFSs to model the linguistic labels of the system. We define the *weak ignorance function*, based on the ignorance function [11] for computing the degree of ignorance (lack of information) associated with the definition of the membership functions. We also propose a new parametrized IVFS construction method. This method allows us to automatically construct an IVFS from a given fuzzy set using weak ignorance functions. In this manner, we can build the IVFSs model from any initial knowledge base in such a way that the length of each IVFS is proportional to the degree of ignorance. The IVFSs model maintains the interpretability of the system as it uses the same linguistic labels and also the same rules. The representation of the linguistic labels by means of IVFSs leads to a natural extension of both the Fuzzy Reasoning Method (FRM) and the computation of the rule weight.

Furthermore, we propose the definition of a single methodology that allows the simultaneous tuning of both the degree of ignorance and the lateral position (based on the 2-tuples fuzzy linguistic representation [31]) of the linguistic labels. This methodology can exploit the good features of both tuning approaches, leading to a better adaptation of the fuzzy partitions. Its final aim is to improve the performance of FRBCSs in a general framework.

In order to show the goodness of our proposed methodology, we will use the Fuzzy Hybrid Genetics-Based Machine Learning (FH-GBML) rule learning algorithm [37] in order to generate the initial knowledge base. This algorithm has been shown to be able to generate a robust FRBCS [25, 41], offering a solid base on which to apply our methodology, and therefore allowing a high classification accuracy to be achieved.

The study is aimed to show whether the results obtained by the application of our methodology enhances the results of the base FH-GBML algorithm. Furthermore, we evaluate the differences in performance with respect to the two tuning approaches that compose our model when they are applied separately. Moreover, we show the good management of the uncertainty of our methodology with respect to the tuning model based on the linguistic 3-tuples representation [3], which performs a tuning of both the amplitude of the support and the lateral position of fuzzy sets but does not take into account the ignorance inherent in the definition of the linguistic labels. To do so, we perform the experimental study over twenty four data-sets selected from the KEEL data-set repository [6, 7] (`http://www.keel.es/dataset.php`). In order to give statistical support to the findings extracted from the experimental analysis, we carry out some non-parametric tests as suggested in the specialized literature [21, 27, 29].

This paper is organized as follows: in Section 2 we introduce some basic concepts of FRBCSs together with the description of the rule learning algorithm considered in this paper. In Section 3 we describe a procedure to build IVFSs starting from fuzzy sets and using weak ignorance functions. Section 4 describes in detail the FRBCS with the linguistic labels modeled by means of IVFSs and the modifications performed in the FRM in order to work with this representation of the linguistic terms. Then, in Section 5 we describe the methodology proposed for the cooperative evolutionary tuning approach. Finally, we show an experimental study in Section 6 and we summarize the paper with the main concluding remarks in Section 7.

2. Fuzzy Rule-Based Classification Systems

In this section we present a brief introduction of FRBCSs. First, we will describe the type of fuzzy rules used in this work, together with the rule weight and the FRM. Next, we will present the FH-GBML rule learning algorithm [37].

2.1. Fuzzy rule-based classification systems

There are a lot of techniques in the Data Mining field to deal with the classification problem. Among them, FRBCSs provide an interpretable model by means of the use of linguistic labels in their rules.

Consider m labeled patterns $x_p = (x_{p1}, \ldots, x_{pn})$, $p = 1, 2, \ldots, m$ where x_{pi} is the ith attribute value ($i = 1, 2, \ldots, n$). We have a set of linguistic values (and their membership functions) describing each attribute. We use fuzzy rules of the following form:

$$\text{Rule } R_j : \text{ If } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \text{ then Class} = C_j \text{ with } RW_j \tag{1}$$

where R_j is the label of the jth rule, $x = (x_1, \ldots, x_n)$ is an n-dimensional pattern vector, A_{ji} is an antecedent fuzzy set representing a linguistic term, C_j is a class label, and RW_j is the rule weight [35]. Specifically, in this paper the rule weight is computed using the Penalized Certainty Factor defined in [38] as:

$$PCF_j = \frac{\sum\limits_{x_p \in Class C_j} \mu_{A_j}(x_p) - \sum\limits_{x_p \notin Class C_j} \mu_{A_j}(x_p)}{\sum\limits_{p=1}^{m} \mu_{A_j}(x_p)} \tag{2}$$

Let $x_p = (x_{p1}, \ldots, x_{pn})$ be a new pattern, L denote the number of rules in the rule base and M the number of classes of the problem; then, the steps of the FRM [16] are as follows:

1. *Matching degree*, that is, *the strength of activation of the if-part for all rules in the rule base with the pattern x_p.* A conjunction operator (t-norm) is applied in order to carry out this computation.

$$\mu_{A_j}(x_p) = T(\mu_{A_{j1}}(x_{p1}), \ldots, \mu_{A_{jn}}(x_{pn})), \qquad j = 1, \ldots, L. \tag{3}$$

2. *Association degree.* To compute the *association degree of the pattern x_p with the M classes according to each rule in the rule base.* When using rules in the form shown in (1) this association degree only refers to the consequent class of the rule (i.e. $k = Class(R_j)$).

$$\mathrm{b}_j^k = h(\mu_{A_j}(x_p), RW_j^k) \qquad k = 1, \ldots, M, \quad j = 1, \ldots, L. \tag{4}$$

3. *Pattern classification soundness degree for all classes.* We use an aggregation function that combines the positive degrees of association calculated in the previous step.

$$Y_k = f(b_j^k,\ j = 1, \ldots, L \text{ and } b_j^k > 0), \qquad k = 1, \ldots, M. \tag{5}$$

4. *Classification.* We apply a decision function F over the soundness degree of the system for the pattern classification for all classes. This function will determine the class label l corresponding to the maximum value.

$$F(Y_1, \ldots, Y_M) = \arg \max_{k=1,\ldots,M} (Y_k) \tag{6}$$

2.2. Fuzzy hybrid genetic based machine learning rule generation algorithm

Different GFSs have been proposed in the specialized literature for designing fuzzy rule-based systems in order to avoid the necessity of linguistic knowledge from domain experts [18, 26, 33, 44].

The basis of the method described here, the FH-GBML algorithm [37], consists of a Pittsburgh approach where each rule set is handled as an individual. It also contains a Genetic Cooperative Competitive Learning (GCCL) approach (an individual represents a unique rule), which is used as a kind of heuristic mutation for partially modifying each rule set, because of its high search ability to efficiently find good fuzzy rules.

This method simultaneously uses four fuzzy set partitions for each attribute, as shown in Figure 1. As a result, each antecedent attribute is initially associated with 14 fuzzy sets generated by these four partitions as well as a special "do not care" set, i.e., 15 in total.

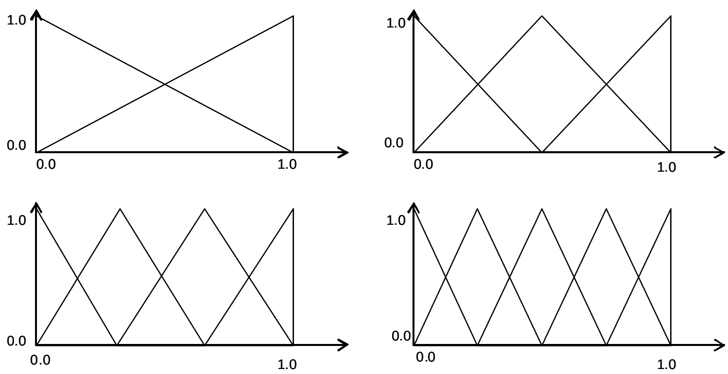


Figure 1: Four fuzzy partitions for each attribute membership function

The main steps of this algorithm are described below:

Step 1: Generate N_{pop} rule sets with N_{rule} fuzzy rules.

Step 2: Calculate the fitness value of each rule set in the current population.

Step 3: Generate (N_{pop} -1) rule sets by selection, crossover and mutation in the same manner as the Pittsburgh-style algorithm. Apply a single iteration of the GCCL-style algorithm (i.e., the rule generation and the replacement) to each of the generated rule sets with a pre-specified probability.

Step 4: Add the best rule set in the current population to the newly generated (N_{pop} -1) rule sets to form the next population.

Step 5: Return to Step 2 if the pre-specified stopping condition is not satisfied.

Next, we will describe every step of the algorithm:

- Initialization: N_{rule} training patterns are randomly selected. Then, a fuzzy rule from each of the selected training patterns is generated by choosing probabilistically (as shown in (7)) an antecedent fuzzy set from the 14 candidates $B_k(k = 1, 2, \ldots, 14)$ (see Figure 1) for each attribute. Then each antecedent fuzzy set of the generated fuzzy rule is replaced with *don't care* using a pre-specified probability $P_{don't\ care}$.

$$P(B_k) = \frac{\mu_{B_k}(x_{pi})}{\sum_{j=1}^{14} \mu_{B_j}(x_{pi})} \tag{7}$$

- Fitness computation: The fitness value of each rule set S_i in the current population is calculated as the number of correctly classified training patterns by S_i. For the GCCL approach the computation follows the same scheme, counting the number of correct hits for each single rule.

- Selection: It is based on binary tournament.

- Crossover: The substring-wise and bit-wise uniform crossover are applied in the Pittsburgh part. In the case of the GCCL part only the bit-wise uniform crossover is considered.

- Mutation: Each fuzzy partition of the individuals is randomly replaced with a different fuzzy partition using a pre-specified mutation probability for both approaches.

For more details about this proposal, please refer to [37].

3. Linguistic Labels Modeled by means of Interval-Valued Fuzzy Sets

In this section, we will recall the definitions of IVFS and ignorance functions [11]. Furthermore, we propose the use of the ignorance function depending on only one variable and we denote this new function as the *weak ignorance function*. Finally, we present a new construction method of IVFSs starting from fuzzy sets and using weak ignorance functions.

3.1. Interval-valued fuzzy sets

In the introduction of this paper, we stressed the significance of selecting the membership functions for modeling the linguistic terms fitted as much as possible for a specific problem and, therefore, allowing a better representation of the knowledge. In particular, the selection of the membership function must be done while keeping as much information as possible. IVFSs allow us to do this because they assign as membership an interval instead of a single number. These sets were born in the 1970s with the work of Sambuc [48]. In the 80s, Gorzalczany denoted these sets for the first time as IVFSs [30].

We should point out that IVFSs are also known as Interval Type-2 Fuzzy Sets in the specialized literature. For example, Liang and Mendel have carried out a deep study about these sets in [40], Wu and Mendel gave uncertainty measures for these sets in [56] and also operations for type-2 fuzzy sets are given in [15, 22, 52, 54].

Definition 1. *An Interval-valued Fuzzy Set (IVFS) A on the universe $U \neq \emptyset$ is given by:*

$$A = \{(u, A(u))|u \in U\}$$

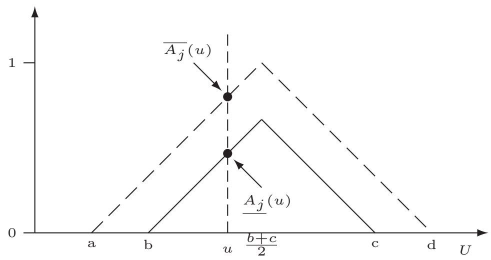
where

$$A(u) = [\underline{A}(u), \overline{A}(u)] \in L([0,1])$$

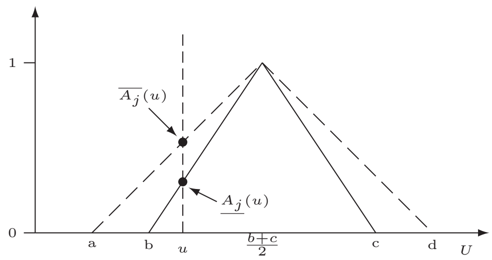
being

$$L([0,1]) = \{\mathbf{x} = [\underline{x}, \overline{x}]|(\underline{x}, \overline{x}) \in [0,1]^2 \text{ and } \underline{x} \leq \overline{x}\}$$

Obviously, $A(u) = [\underline{A}(u), \overline{A}(u)]$ is the membership degree of $u \in U$. Figure 2 depicts two examples of IVFSs: the interval $[\underline{A_j}(u), \overline{A_j}(u)]$ and not a number from $[0,1]$ is assigned as the membership to each element $u \in U$.



(a) Generic IVFS



(b) IVFS employed in this work

Figure 2: Examples of IVFSs. The points a, b, $\frac{b+c}{2}$, c, d define the bounds of the IVFS.

3.2. Weak ignorance function

The concept of ignorance functions is defined in [11] in order to quantify the uncertainty that an expert has when he or she assigns numerical values to the memberships of an element to two different classes.

Definition 2. *An ignorance function is a continuous mapping* $G_i : [0,1]^2 \rightarrow [0,1]$ *so that:*

(G_i1) $G_i(x,y) = G_i(y,x)$ *for all* $x, y \in [0,1]$;

(G_i2) $G_i(x,y) = 0$ *if and only if* $x = 1$ *or* $y = 1$;

(G_i3) $If x = 0.5$ *and* $y = 0.5$, *then* $G_i(x,y) = 1$;

(G_i4) G_i *is decreasing in* $[0.5,1]^2$;

(G_i5) G_i *is increasing in* $[0,0.5]^2$.

Observe that this definition implies that we have assumed that a value of 0.5 corresponds to the complete lack of knowledge of the expert on the membership of an element to a class. In [11] the five axioms in Definition 2 are justified.

Proposition 1. *Let* $G_i : [0,1]^2 \rightarrow [0,1]$ *be an ignorance function. The mapping:*

$$g : [0,1] \rightarrow [0,1] \text{ given by}$$

$$g(x) = G_i(x, 1-x)$$

is a continuous function that satisfies:

(i) $g(x) = g(1-x)$ *for all* $x \in [0,1]$;

(ii) $g(x) = 0$ *if and only if* $x = 0$ *or* $x = 1$;

(iii) $g(0.5) = 1$.

Proof 1. *It is enough to take into account Definition 2.* □

Definition 3. *A continuous mapping* $g : [0,1] \rightarrow [0,1]$ *is called weak ignorance function if it satisfies the items (i) - (iii) in Proposition 1.*

The name is due to the fact that they are only associated with one element, in the sense that they depend on a single variable, and not on two. We will use weak ignorance functions in order to quantify the lack of knowledge that exists when assigning a numerical value to the membership of an element to a given linguistic label.

Example 1. *The following function:* $g(x) = 2\min(x, 1-x)$, *is a weak ignorance function.*

3.3. Construction of interval-valued fuzzy sets starting from fuzzy sets and weak ignorance functions

In this paper, we use bijections in order to construct IVFSs in such a way that their amplitude is proportional to the weak ignorance function having the maximum ignorance degree when the initial membership degree is 0. Therefore, we employ a bijection $h : [0,1] \to [0.5,1]$ so that $h(0) = 0.5$ and $h(1) = 1$. Specifically, we employ the following one:

$$h(x) = \frac{1}{2}x + \frac{1}{2}$$

We build the upper bound of an IVFS (dashed line in Figure 2(b)) starting from a fuzzy set, like the one depicted with a solid line in Figure 2(b) where the membership function of this fuzzy set is:

$$\mu_A(x) = \begin{cases} 0, & \text{if } x < b, \\ \frac{2}{c-b}(x-b), & \text{if } b \leq x \leq \frac{b+c}{2} \text{ and } c \neq b, \\ \frac{2}{b-c}(x-c), & \text{if } \frac{b+c}{2} < x \leq c \text{ and } c \neq b, \\ 0, & \text{if } c < x, \end{cases} \tag{8}$$

We must point out that we call r_1 the straight line that joins the points $(b,0)$ and $\left(\frac{b+c}{2}, 1\right)$.

$$r_1 \equiv y = \frac{2}{c-b}(x-b) \tag{9}$$

By symmetry, the study is only carried out for the left side of the triangular membership function, that is, for the values less than or equal to $\frac{b+c}{2}$.

We can define the point a in function of the points of the previous straight line so that $a = b - W(c-b)$, where $W \geq 0$ stands for the displacement with respect to the point b. Considering this fact, the equation of the straight line that joins the points $(a,0)$ and $\left(\frac{b+c}{2}, 1\right)$ is:

$$r_2 \equiv y = \frac{2}{(c-b)(1+2W)}(x - b + W(c-b)) \tag{10}$$

Given the functions h and g previously defined, for simplicity we denote $H = h \circ g$.

In the following proposition we present the construction method of IVFSs from a fuzzy set and the weak ignorance function associated with this fuzzy set:

Proposition 2. *Let the fuzzy set A given in Eq. (8) and let $\beta(x), \gamma(x) \in [0,1]$. Under these conditions:*

$$\left[\mu_A(x), \mu_A(x)\big(1 - \beta(x)H\big(\mu_A(x)\big)\big) + \gamma(x)H\big(\mu_A(x)\big)\right] \in L([0,1])$$

if and only if

$$\beta(x)\mu_A(x) \leq \gamma(x) \leq \frac{1}{H(\mu_A(x))}\big(1 - \mu_A(x)\big(1 - \beta(x)H(\mu_A(x))\big)\big) \qquad \textit{for all } x \in \Re$$

Proof 2. *Direct.* □

Remark 1. *Note that under the conditions of Proposition 2:* $H(\mu_A(x)) \neq 0 \qquad$ *for all* $x \in \Re$

We must point out that $\mu_A(x)$ is $r_1(x)$ between (b, 0) and ($\frac{b+c}{2}$,1) and therefore, the expression of the upper bound given in Proposition 2 is $r_2(x)$ between (a, 0) and ($\frac{b+c}{2}$,1).

It is easy to prove that in the construction method presented in Proposition 2 the length of the intervals, L, is always proportional to the weak ignorance function up to a bijection. In fact, we have the following result:

Proposition 3. *The length of the interval built in Proposition 2 is:*

$$L(x) = \big(\gamma(x) - \mu_A(x)\beta(x)\big)H(\mu_A(x)) \qquad \textit{for all } x \in \Re$$

Proof 3. *Direct.* □

4. Fuzzy Rule-Based Classification Systems based on Interval-Valued Fuzzy Sets

As we have stated in the introduction section, we propose to model the linguistic labels by means of IVFSs in order to take into account the uncertainty inherent in the definition of the membership functions. This modeling implies that the rule weight must be composed of two numbers and we also have to adapt the FRM to work with this representation.

According to these issues, the construction of the IVFS partition is explained in Subsection 4.1 and both the extension of the FRM and the computation of the rule weight are described in Subsection 4.2.

4.1. Interval-valued fuzzy sets based fuzzy partition

In this paper we employ fuzzy partitions composed of IVFSs with the structure that has been depicted in Figure 2(b) (Section 3), where the solid line represents the lower bound ($\underline{A_j}$) and the dashed line represents the upper bound ($\overline{A_j}$).

The construction method of each IVFS is the one presented in Subsection 3.3 considering as the initial fuzzy sets those defined by expert knowledge or defined homogeneously over the input space. We must point out that the parameter W can take values in the interval $[0, 0.5]$. The upper and the lower bounds of this interval correspond to the following notable situations:

- Total certainty ($W = 0$): the linguistic label is perfectly defined by means of the initial fuzzy set. That is, when the lower and the upper bounds of an IVFS are the same.
- Maximum uncertainty ($W = 0.5$): there is a huge lack of information in the definition of the linguistic label. So, the amplitude of the upper bound of the IVFS is twice the amplitude of the lower bound.

Therefore, in the initial construction we set the parameter W to 0.25 since it is the intermediate situation between the ones previously described. In this manner, the amplitude of the upper bound is 50% greater than the one of the lower bound (details are given in Subsection 5.1).

We will employ rules in the form presented in Subsection 2.1 where each A_{ji} will be an IVFS instead of a fuzzy set.

4.2. Fuzzy reasoning method

The modifications in the structure of the fuzzy labels also imply an extension of the original FRM used for classifying new patterns. Let $x_p = (x_{p1}, \ldots, x_{pn})$ be a new pattern; the general model of fuzzy reasoning for classification, presented in Subsection 2.1, will be modified in the following way:

1. We have two *matching degrees* because we are working with an interval, each one will be associated with the lower and the upper bound respectively and will be calculated applying a T-norm in the following way:

$$\mu_L A_j(x_p) = T(\underline{A_{j1}}(x_{p1}), \ldots, \underline{A_{jn}}(x_{pn})), \qquad j = 1, \ldots, L. \tag{11}$$

$$\mu_U A_j(x_p) = T(\overline{A_{j1}}(x_{p1}), \ldots, \overline{A_{jn}}(x_{pn})), \qquad j = 1, \ldots, L. \tag{12}$$

   We apply a T-norm to both lower and upper bounds. Therefore, the matching degrees obtained form the following interval: $[\mu_L A_j(x_p), \mu_U A_j(x_p)]$.
2. As *association degree* we take the mean of the product of the matching degree by the rule weight, which is composed of two numbers, associated with the lower and the upper bound respectively. The rule weights will be denoted as PCF_{Lj} and PCF_{Uj} and their computation will be made following Expression (2), considering the lower and the upper bounds to be the terms in each case, that is:

$$PCF_{Lj} = \frac{\displaystyle\sum_{x_p \in Class C_j} \underline{A_j}(x_p) - \sum_{x_p \notin Class C_j} \underline{A_j}(x_p)}{\displaystyle\sum_{p=1}^{m} \underline{A_j}(x_p)} \tag{13}$$

$$PCF_{Uj} = \frac{\displaystyle\sum_{x_p \in Class C_j} \overline{A_j}(x_p) - \sum_{x_p \notin Class C_j} \overline{A_j}(x_p)}{\displaystyle\sum_{p=1}^{m} \overline{A_j}(x_p)} \tag{14}$$

   The final association degree will be computed as follows:

$$b_j^k = \frac{\mu_L A_j(x_p) * PCF_L j^k + \mu_U A_j(x_p) * PCF_U j^k}{2}$$
$$k = 1, \ldots, M, \quad j = 1, \ldots, L. \tag{15}$$

At this point we already have a single value associated with the class. Accordingly, we can apply the rest of the method in the same way as in the general FRM presented in [16].

5. Genetic Tuning of the Data Base for Interval-Valued Fuzzy Sets

Most of the FRBCSs proposed in the literature, in particular the FH-GBML algorithm, are mainly focused on determining the set of fuzzy rules composing the rule base of the model, rather than finding an optimal definition of the data base, which may cause the cooperative behavior of the rules to be less than optimal. Specifically, the membership functions that compose the data base are usually obtained by a normalization process and they remain fixed during the rule base derivation process. This fact usually implies that the membership functions are not properly adapted to the context of each variable, which limits the performance of the global rule set.

To solve this problem, a post-processing tuning step is often used. This methodology consists of refining the initial definition of the data base once the rule base has been obtained [2, 4, 39]. The tuning introduces a variation in the shape of the membership functions that improves their global interaction with the main aim of inducing better cooperation among the rules [5, 12]. In this way, the real aim of the tuning is to find the best global configuration of the membership functions and not only to find specific membership functions in an independent way.

In order to carry out this tuning step, in our previous work on the topic we considered tuning the amplitude of the support of the upper bound [49] and in this way to manage the ignorance that each IVFS represents. Also, in the specialized literature we found an approach to make a lateral adjustment of the fuzzy labels based on the 2-tuples fuzzy linguistic representation which obtained very accurate models [5, 25]. In this work we want to go one step further and combine these two approaches in a cooperative model.

In the remainder of this section we will first introduce the two tuning approaches and then we will present our proposal to carry out both tunings simultaneously. Finally, we will describe in detail the evolutionary process for tuning the membership functions.

5.1. Weak ignorance tuning of the linguistic labels

In Subsection 4.1 we have shown that the initial construction of the IVFSs is performed by associating an intermediate degree of uncertainty with the definition of the linguistic labels. In this manner, the amplitude of the support of the upper bound is fixed for all the IVFSs but, as the amount of uncertainty associated with the definition of each linguistic label can differ, the support of the upper bound of each IVFS is variable. In other words, the degree of ignorance each IVFS represents does not need to be the same for all IVFSs in the data base [49].

To deal with this problem, we apply a post-processing genetic tuning step in order to improve the behavior of the FRBCSs performing slight changes to the amplitude of the upper bound of the IVFSs. In this way, by varying the amplitude of the upper bound of the IVFS we vary the degree of ignorance for all the elements in the support of the linguistic label.

As we have stated in Subsection 4.1, we can modify the amplitude of the upper bound of the IVFSs according to the value of the parameter W. The interval in which W can vary is [0, 0.5], that is, from the situation without ignorance (W = 0) to the situation of maximum ignorance allowed in our model (W = 0.5). The degree of ignorance that each IVFS represents will be uniformly increased according to intermediate values being W = 0.25 the initial situation as we have explained in Section 4. These situations are depicted in Figure 3.

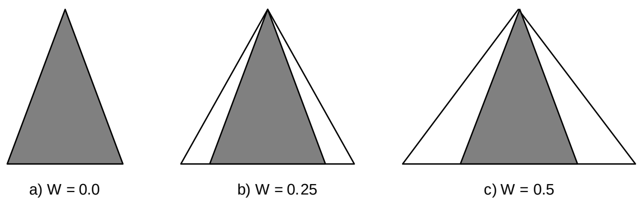


Figure 3: Parameter values representation in the genetic weak ignorance tuning. a) No ignorance. b) Initial situation. c) Maximum ignorance.

5.2. Lateral tuning of linguistic labels based on the 2-tuples fuzzy linguistic model

In our initial model we have fixed the position of the different labels in such a way that for each value of the input space of each variable the sum of the membership degrees of the different labels is 1. For example, any of the four fuzzy partitions depicted in Figure 1 fulfills this condition. This label distribution does not need to be optimal as the data distribution is not usually uniform. Therefore, we use the genetic tuning based on the 2-tuples fuzzy linguistic model [31] to make the lateral displacements of the linguistic labels.

The symbolic translation of a linguistic term which is given by a number within the interval [-0.5, 0.5). This number expresses the bounds of the domain of a label when it is moving between its two lateral labels. Let us consider a set of labels S representing a fuzzy partition. Formally, we have the pair, $(s_i, \alpha_i), s_i \in S, \alpha_i \in [-0.5, 0.5)$. An example is illustrated in Figure 4, where we show the symbolic translation of a label represented by the pair (S_2,-0.3) together with the lateral displacement of the corresponding membership function.

We must point out that we use the Global Tuning of the Semantics method[5]. In this model, the tuning is applied at the data base level, maintaining the interpretability of the full system to a large degree. In other words, the pair (X_i, label) will take the same tuning value in all the rules where the linguistic partition *label* is considered for the attribute X_i. For example, X_i is $(\mathit{Low}, 0.15)$ will present the same value for those rules in which the couple "X_i is *Low*" is initially considered. Therefore, only one displacement parameter is considered for each label in the data base.

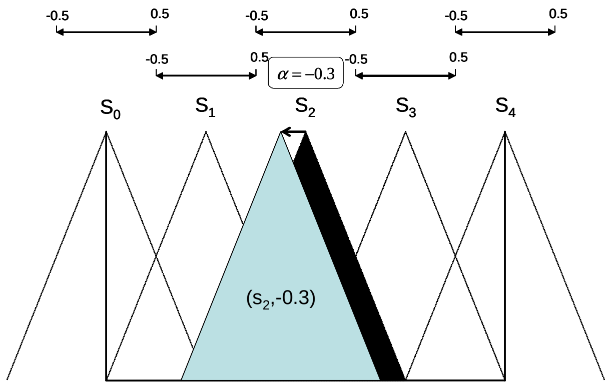


Figure 4: Lateral displacement of a membership function.

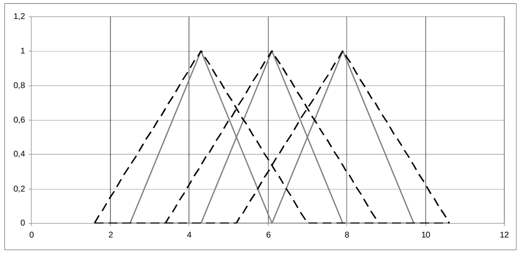
5.3. Cooperative tuning for both the weak ignorance and the lateral displacement

In this paper our aim is to achieve a positive synergy between the tuning of the weak ignorance and the lateral tuning. The motivation for this proposal lies in the following issue: the isolated use of each tuning approach might cause the system to reach a sub-optimal model. With this cooperative methodology, the GA search engine will be able to work at the same time with both characteristics, and this should lead us to discover solutions better adapted to the problem and therefore more accurate ones.

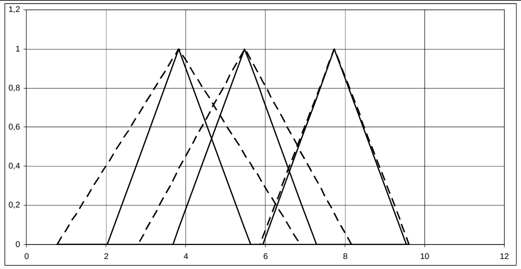
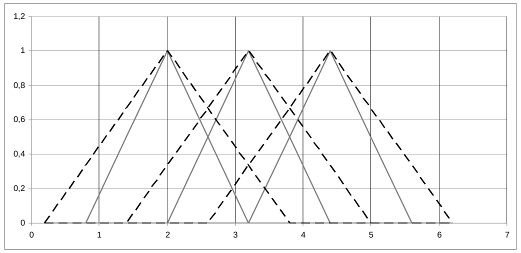
To do this, it is enough to define a unique representation inside the GA for both possibilities. However, we have to consider the size of the search space, since a high number of parameters can lead to poor solutions (e.g. in the case of performing the tuning taking into account the five parameters depicted in Figure 2(b) that define the bounds of the IVFS). Therefore, we propose to downsize the search space by decreasing the number of parameters (see Subsection 5.4.1 for details about the representation) in the following way:

1. To perform the lateral displacement we consider one parameter representing each linguistic label within each variable.
2. To perform the weak ignorance tuning, we consider only one parameter for each linguistic label.

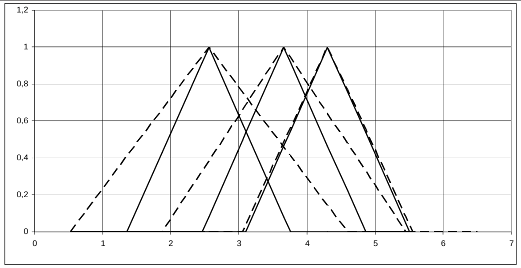
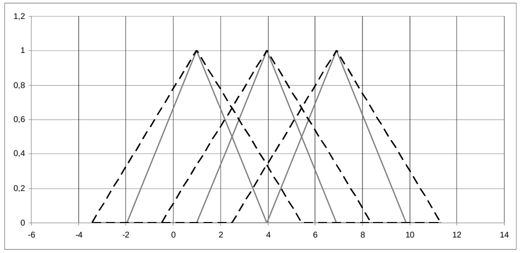
In order to illustrate the effect of the cooperative evolutionary tuning model in the fuzzy partitions, Table 1 depicts the results of the tuning process for the *iris* data-set in columns. The first column refers to the variable studied, the second column shows the fuzzy partition in the initial FRBCS with IVFSs generated by the FH-GBML algorithm, the third presents the final values of the genes after the tuning process and the last column depicts the representation of each fuzzy partition after the post-processing step. We must point out that, for the sake of simplicity, from the four possible fuzzy partitions that the FH-GBML algorithm contemplates, we only depict the one corresponding to a

<table>
<tr><th></th><th>Initial partition</th><th>Gene values for each linguistic label (LL)</th><th>Final partition</th></tr>
<tr><td>Sepal length</td><td></td><td>

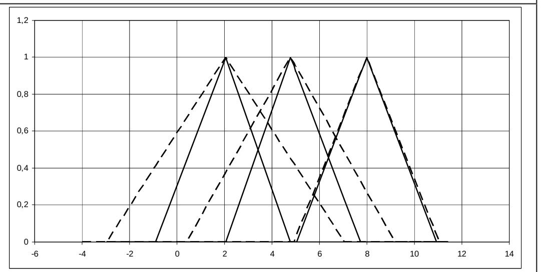
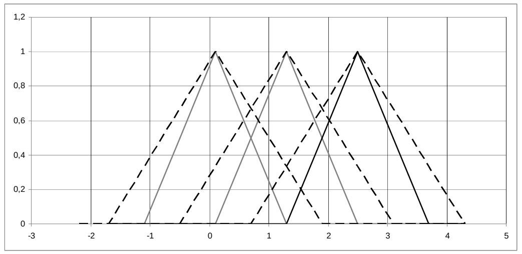
Tuning	LL. 1	LL. 2	LL. 3
Lat	-0.26	-0.34	-0.09
WI	0.34	0.24	0.02

</td><td></td></tr>
<tr><td>Sepal width</td><td></td><td>

Tuning	LL. 1	LL. 2	LL. 3
Lat	0.46	0.38	-0.08
WI	0.34	0.24	0.02

</td><td></td></tr>
<tr><td>Petal length</td><td></td><td>

Tuning	LL. 1	LL. 2	LL. 3
Lat	0.35	0.28	0.37
WI	0.34	0.24	0.02

</td><td></td></tr>
<tr><td>Petal width</td><td></td><td>

Tuning	LL. 1	LL. 2	LL. 3
Lat	-0.23	-0.27	0.09
WI	0.34	0.24	0.02

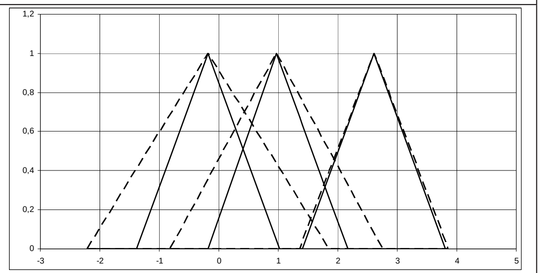
</td><td></td></tr>
</table>

Table 1: Fuzzy partitions of the *iris* data-set before and after of the cooperative evolutionary tuning model.

degree of granularity of three labels per variable. First of all, the final values of the genes confirm the necessity of the contextualization of each linguistic label as no fuzzy partition remains in its initial state. Furthermore, we show the good choice of the third linguistic label, since after the application of the cooperative evolutionary tuning model the uncertainty that the IVFS represents is practically nil and most of the variables only have a slight lateral displacement. We observe that the most processed variable is the *petal length*; this fact shows the goodness of the solution as, knowing the features of this data-set, this variable together with the *petal width* are enough to discriminate well among all the classes.

5.4. Evolutionary tuning

We must recall from the introduction to the paper that our aim is to use a hybridization of FRBCSs and GAs in order to look for the optimal configuration of the parameters of the knowledge base. From the optimization point of view, to find an appropriate fuzzy model is equivalent to coding it as a parameter structure and then finding the parameter values that give us the optimum for a specific fitness function.

In the remainder of this subsection, we introduce the features depending on the specific approach and then we describe the evolutionary method employed in this paper.

5.4.1. Representation and fitness function

The tuning of the membership functions' parameters can be considered as a search problem to which GAs can be applied. To accomplish this, we take into account two important issues: the specification of the representation of the solutions and the definition of the fitness function.

- *Representation*: In this paper we employ three different tuning approaches, introduced in the previous subsections, where a real coding is considered in all models. However, the representation of each gene of the chromosome depends on the approach. Let us consider the following number of labels per variable: $(m^1, m^2, ..., m^n)$, with n being the number of variables. Then, the representations of the chromosome are as follows:
  - *Weak ignorance tuning:* Each gene represents the modification of the degree of ignorance of the linguistic labels in the data base as we have explained in Subsection 5.1. The structure of the chromosome is:

    $$C_{WI} = (a_{11}, ..., a_{1m^1}, a_{21}, ..., a_{2m^2}, ..., a_{n1}, ..., a_{nm^n})$$

    Then, the chromosome length is equal to the number of labels times the number of variables.
  - *Lateral tuning:* Each gene represents the lateral displacement of the linguistic labels in the data base as we have explained in Subsec-

tion 5.2. The structure of the chromosome is:

$$C_{Lat} = (l_{11}, ..., l_{1m^1}, l_{21}, ..., l_{2m^2}, ..., l_{n1}, ..., l_{nm^n})$$

Hence, the chromosome length is equal to the number of labels times the number of variables.

  – *Cooperative tuning:* Each chromosome will be composed of two parts, one to perform the weak ignorance tuning and the other to perform the lateral tuning. Then, a chromosome has the structure:

$$C_{Coop} = (C_{Lat} + a_1, a_2, ..., a_{14})$$

Therefore, the chromosome length is equal to the number of variables times the number of linguistic labels plus the number of linguistic labels.

- *Fitness function*: We employ the most common metric for classification, i.e. the accuracy rate.

The generic code structure and independent performance features of GAs make them suitable candidates to incorporate a priori knowledge. In the case of FRBSs, this a priori knowledge may be in the form of linguistic variables, fuzzy membership function parameters, fuzzy rules, number of rules, etc. Therefore, we initialize the population of the different approaches considered in this paper in the following way:

- *Weak ignorance tuning:* The initial pool is obtained with the first individual having all genes with a value of 0.25 (the initial FRBCS). The second and the third individuals have all genes with values of 0 and 0.5 respectively, whereas the remaining individuals are generated at random in [0, 0.5].

- *Lateral tuning:* The initial pool is obtained with the first individual having all genes with a value of 0.0 (the initial FRBCS), whereas the remaining individuals are generated at random in [-0.5, 0.5).

- *Cooperative tuning:* In this model we initialize three individuals having all genes employed to perform the lateral tuning with a value of 0.0, whereas the genes used to carry out the weak ignorance tuning have values of 0, 0.25 and 0.5 respectively. The remaining individuals will have initialized all the genes randomly.

5.4.2. Evolutionary model

CHC[23] is a classical evolutionary model that introduces different features to obtain a trade-off between exploration and exploitation; such as incest prevention, reinitialization of the search process when it becomes blocked and the competition between parents and offspring in the replacement process. In accordance with both these suitable features and with the models used in [5, 25], we use CHC to deal with the evolutionary tuning.

During each generation the CHC algorithm develops the following steps.

- It uses a parent population of size N to generate an intermediate population of N individuals that are randomly paired and used to generate N potential offspring.
- A survival competition is held where the best N chromosomes from the parent and offspring populations are selected to form the next generation.

In order to perform the crossover operator, we consider the Parent Centric BLX (PCBLX) operator [32], which is based on the BLX-α. Figure 5 depicts the behavior of these kinds of operators.

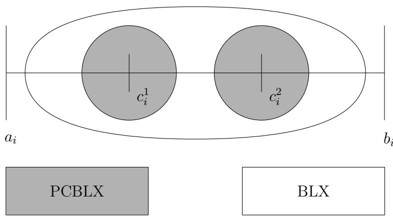


Figure 5: Scheme of the behavior of the BLX and PCBLX operators

PCBLX is described as follows. Let us assume that $X = (x_1 \cdots x_n)$ and $Y = (y_1 \cdots y_n)$, $(x_i, y_i \in [a_i, b_i] \subset \Re, i = 1 \cdots n)$, are two real-coded chromosomes that are going to be crossed. PCBLX operator generates the two following offspring:

- $O_1 = (o_{11} \cdots o_{1n})$, where o_{1i} is a randomly (uniformly) chosen number from the interval $[l_i^1, u_i^1]$, with $l_i^1 = max\{a_i, x_i - I_i\}$, $u_i^1 = min\{b_i, x_i + I_i\}$, and $I_i =\mid x_i - y_i \mid$.
- $O_2 = (o_{21} \cdots o_{2n})$, where o_{2i} is a randomly (uniformly) chosen number from the interval $[l_i^2, u_i^2]$, with $l_i^2 = max\{a_i, y_i - I_i\}$ and $u_i^2 = min\{b_i, y_i + I_i\}$.

On the other hand, the incest prevention mechanism will only be considered in order to apply the PCBLX operator. In our case, two parents are crossed if half their Hamming distance is above a predetermined threshold, L. Since we consider a real coding scheme, we have to transform each gene considering a Gray Code (binary code) with a fixed number of bits per gene ($BITSGENE$), which is determined by the system expert. In this way, the threshold value is initialized as:

$$L = (\#Genes \cdot BITSGENE)/4.0$$

where $\#Genes$ stands for the total length of the chromosome. Following the original CHC scheme, L is decremented by one ($BITSGENE$ in this case) when there are no new individuals in the next generation.

In order to work with the cooperative evolutionary tuning model, we make the crossover in the following way:

- We cross the parts of the chromosome representing the same kind of tuning among themselves. In this process we generate four offspring: they are all the possible combinations of the two offspring generated by the crossover of the lateral tuning part and two more created by the crossover of the weak ignorance tuning part.
- We select the two best ones to be included in the population.

No mutation is applied during the recombination phase. Instead, when the threshold value is lower than zero, all the chromosomes are regenerated randomly to introduce new diversity to the search. The best global solution found is included in the population to increase the convergence of the algorithm as in the elitist scheme.

6. Experimental Study

In this study, our aim is to show the improvement achieved in FRBCSs by means of the combination of the IVFSs model and the cooperative evolutionary tuning model. Specifically, we check the usefulness of our methodology by means of comparisons with respect to the achieved results by the initial FH-GBML FRBCS with and without lateral tuning. In addition, we study the behavior of the cooperative evolutionary tuning methodology with respect to the two tuning approaches when they are applied separately over the FRBCS with the linguistic labels modeled by means of IVFSs.

Moreover, we also include in the study the tuning approach based on the linguistic 3-tuples representation [3]. This approach performs both a lateral displacement and a modification of the amplitude of the support of the fuzzy sets but without taking into account the uncertainty inherent in the definition of the linguistic labels. The aim of this comparison is to show the good management of the uncertainty of our cooperative evolutionary tuning model leading to a global performance improvement.

In this section, first we describe the experimental set-up together with the parameters employed in the study and the statistical test considered in this paper. Next, we introduce the experimental study carried out.

6.1. Experimental set-up

We have selected a wide benchmark of 24 data-sets selected from the KEEL data-set repository [6, 7] (`http://www.keel.es/dataset.php`). Table 2 summarizes the properties of the selected data-sets, showing for each data-set the number of examples (#Ex.), the number of attributes (#Atts.), the number of numerical (#Num.) and nominal (#Nom.) features, and the number of classes (#Class.). The *magic, page-blocks, penbased* and *ring* data-sets have been stratified sampled at 10% in order to reduce their size for training. In the case of missing values (*autos, cleveland, crx, hepatitis* and *wisconsin*) we have removed those instances from the data-set.

Table 2: Summary Description for the employed data-sets.

Id.	**Data-set**	**#Ex.**	**#Atts.**	**#Num.**	**#Nom.**	**#Class.**
aus	Australian	690	14	8	6	2
aut	Autos	150	25	15	10	6
bal	Balance	625	4	4	0	3
bup	Bupa	345	6	6	0	2
cle	Cleveland	297	13	13	0	5
con	Contraceptive	1,473	9	6	3	3
crx	Crx	125	15	6	9	2
eco	Ecoli	336	7	7	0	8
ger	German	1,000	20	6	14	2
gla	Glass	214	9	9	0	6
hab	Haberman	306	3	3	0	2
hea	Heart	270	13	6	7	2
hep	Hepatitis	80	19	6	13	2
iri	Iris	150	4	4	0	3
mag	Magic	1,902	10	10	0	2
new	New-Thyroid	215	5	5	0	3
pag	Page-blocks	548	10	10	0	5
pen	Penbased	1,099	16	16	0	10
pim	Pima	768	8	8	0	2
rin	Ring	740	20	20	0	2
tae	Tae	151	5	3	2	3
veh	Vehicle	846	18	18	0	4
win	Wine	178	13	13	0	3
wis	Wisconsin	683	9	9	0	2

To carry out the different experiments we consider a *5-folder cross-validation model*, i.e., 5 random partitions of data (20% of the patterns), and the combination of 4 of them (80%) as training and the remainder as test. The process was repeated 3 times using different seeds. Therefore, for each data-set we consider the average result over 15 runs.

The selected configuration for the FH-GBML approach consists of product T-norm as conjunction operator, together with the Penalized Certainty Factor approach for the rule weight and FRM of the winning rule. Regarding the specific parameters for the genetic process, we have chosen the following values:

- Number of fuzzy rules: $5 \cdot d$ rules, being 50 the maximum number of fuzzy rules.
- Number of rule sets: 200 rule sets.
- Crossover probability: 0.9.
- Mutation probability: $1/d$.
- Number of replaced rules: All rules except the best-one (Pittsburgh-part, elitist approach), number of rules / 5 (GCCL-part).
- Total number of generations: 1,000 generations.
- Don't care probability: 0.5.
- Probability of the application of the GCCL iteration: 0.5.

where d stands for the dimensionality of the problem (number of variables).

Next, we indicate the values that have been considered for the parameters of the genetic tuning:

- Population Size: 50 individuals.
- Number of evaluations: $5{,}000 \cdot$ d.
- Bits per gene for the Gray codification (for incest prevention): 30 bits.

Finally, we must point out that statistical analysis needs to be carried out in order to find significant differences among the results [28]. We consider the use of non-parametric tests, according to the recommendations made in [21, 27] where a set of simple, safe and robust non-parametric tests for statistical comparisons of classifiers is presented. In this empirical study we will apply pairwise comparisons between the algorithms using the Wilcoxon Signed-Ranks test [51] computing the p-value associated with each comparison, which represents the lowest level of significance of a hypothesis that results in a rejection. In this manner, we can know whether two algorithms are significantly different and how different they are.

6.2. Analysis of the usefulness of the cooperative evolutionary tuning methodology based on IVFSs

Table 3 shows the results achieved by the different approaches, both in train and in test in each data-set. This table presents two different groups of results. The first one covers the approaches in which the data base is composed of standard fuzzy sets. They are:

- The basic FH-GBML method (Base).
- The initial FH-GBML FRBCS postprocessed with lateral tuning (Lat).
- The initial FH-GBML FRBCS postprocessed with the tuning approach based on the linguistic 3-tuples representation (3-tuples).

The second group is formed of the approaches in which the data base is composed of IVFSs. They are:

- The weak ignorance tuning (IVFS_WI).
- The lateral displacement tuning (IVFS_Lat).
- The cooperative evolutionary tuning (IVFS_Coop).

To start with, we analyze the results of our methodology with respect to the results of the approaches belonging to the first group. It is noteworthy that our methodology achieves a high global improvement with respect to the basic FH-GBML FRBCS, obtaining a higher accuracy value in test in twenty out of the twenty four data-sets. Regarding the FH-GBML method using lateral tuning, we observe that our cooperative evolutionary tuning model has a better mean performance, increasing the classification accuracy in sixteen data-sets.

Table 3: Results in Train (Tr.) and Test (Tst) achieved by the different approaches.

Data Set	Base		Lat		3-tuples		IVFS_WI		IVFS_Lat		IVFS_Coop	
	Tr.	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*
aus	87.19	84.78	89.76	85.36	90.09	**86.09**	88.42	85.51	89.66	84.64	89.58	85.36
aut	61.02	45.34	80.82	54.07	80.35	57.26	75.92	54.74	80.03	58.49	80.82	**62.90**
bal	79.04	77.60	84.97	80.16	85.33	**81.12**	80.84	80.48	84.05	80.16	85.21	80.64
bup	70.62	63.48	80.07	64.06	80.29	63.77	74.33	62.61	77.89	**66.67**	78.98	65.51
cle	61.71	54.55	69.49	56.24	71.43	55.23	65.26	**56.91**	65.76	55.56	69.82	55.90
con	47.89	45.22	55.21	50.17	56.07	**51.67**	51.32	48.27	54.70	50.10	56.21	50.72
crx	88.30	86.68	89.38	86.53	89.49	86.53	88.72	**87.29**	89.30	87.14	89.38	86.83
eco	76.33	72.91	83.57	76.19	84.32	76.78	81.26	72.91	84.47	77.98	85.59	**79.76**
ger	73.82	71.30	76.62	71.70	77.20	71.80	75.04	71.60	75.72	**72.30**	77.07	72.10
gla	69.10	57.49	77.68	61.20	77.32	60.27	74.27	57.94	76.85	59.34	77.79	**61.21**
hab	78.26	71.89	80.39	71.89	80.80	**72.23**	78.75	72.22	78.67	72.22	80.39	70.59
hea	85.02	80.37	88.84	**82.22**	89.12	80.37	86.98	81.48	89.02	78.15	89.12	80.00
hep	91.75	81.25	92.70	81.25	92.70	**85.00**	92.06	83.75	92.70	**85.00**	92.70	**85.00**
iri	98.49	95.33	99.16	**97.33**	99.16	96.00	98.82	96.00	98.82	96.00	99.33	96.00
mag	79.02	78.49	82.23	78.81	83.01	79.76	81.19	79.18	82.70	79.92	83.15	**80.18**
new	95.32	92.56	98.71	91.63	99.30	94.42	97.66	93.49	98.95	94.88	99.53	**95.81**
pag	95.88	94.16	96.43	**94.34**	96.52	93.97	96.07	94.16	95.34	93.79	96.71	93.43
pen	69.83	67.18	84.73	76.36	87.21	80.91	83.85	78.27	86.62	81.00	88.62	**83.00**
pim	76.88	73.96	80.31	75.26	80.40	74.74	78.71	75.00	79.65	**75.91**	80.44	75.65
rin	84.97	82.84	92.52	87.30	93.37	86.35	88.43	83.11	92.35	85.54	93.30	**87.43**
tae	62.61	49.01	69.12	53.68	70.62	**54.28**	66.11	52.32	69.11	51.72	70.62	52.99
veh	61.47	58.52	75.94	64.78	76.83	66.43	69.46	62.30	75.55	65.49	75.50	**67.15**
win	97.45	90.97	99.58	91.52	99.86	90.40	98.87	90.97	99.72	**93.79**	99.72	93.76
wis	97.58	95.61	98.09	95.31	98.31	95.17	97.65	**95.75**	98.02	**95.75**	98.28	94.88
Mean	78.73	73.81	84.43	76.14	84.96	76.69	82.08	75.68	83.99	76.73	84.91	**77.37**

Finally, we observe that our methodology improves the performance of the 3-tuples tuning model in fourteen out of the twenty four data-sets. These findings are statistically confirmed in Table 4 where we carry out a pairwise comparison (applying a Wilcoxon test). Here, we observe a higher sum of ranks and a low p-value in all cases, which allow us to determine the enhancement of our approach with the support of statistical differences in favour of our methodology with a high level of confidence.

Table 4: Wilcoxon Test to compare the cooperative evolutionary tuning model (R^+) with the basic FH-GBML with and without tuning (R^-).

Comparison	R^+	R^-	p-value
IVFS_Coop vs. Base	279	21	0.000
IVFS_Coop vs. Lat	219.5	80.5	0.048
IVFS_Coop vs. 3-tuples	209.5	90.5	0.082

The analysis of our methodology with respect to the tuning approaches applied to the FRBCSs with IVFSs yields similar results. The cooperative evolutionary tuning model obtains the best mean performance and, in particular, improves the accuracy in more than half of the data-sets, i.e. sixteen in the case of the weak ignorance tuning and fourteen in the case of the lateral tuning. The statistical study shown in Table 5 (also carried out with a Wilcoxon test) shows that our methodology outperforms both tuning approaches applied to the FH-GBML model with IVFSs with a low p-value.

These results allow us to stress the goodness of our cooperative evolutionary tuning model since it enhances the results provided by both the classical FH-GBML approach and the tuning approaches that compose our model when they are applied separately. Moreover, our proposed methodology enhances the re-

Table 5: Wilcoxon Test to compare the cooperative evolutionary tuning model (R^+) with the tuning approaches over the FRBCS with IVFSs (R^-).

Comparison	R^+	R^-	p-value
IVFS_Coop vs. IVFS_WI	238.5	61.5	0.011
IVFS_Coop vs. IVFS_Lat	225.5	74.5	0.024

sults achieved by the 3-tuples tuning model, an approach which is conceptually similar, since it tunes both the amplitude of the support and the lateral position of fuzzy sets while, however, ignoring the uncertainty inherent in the definition of the membership functions. The improvement provided by our methodology with respect to the remaining ones proves that the good management of the uncertainty associated with the definition of the fuzzy terms has a positive synergy with the lateral tuning, leading to an improvement in the global behavior of the system.

7. Concluding Remarks

In this paper we have introduced two main novelties. On the one hand, we have defined a new construction method of IVFSs starting from fuzzy sets and using weak ignorance functions. This construction method allows us to parametrize the amplitude of the upper bound of each IVFS in such a way that we can analyze the most appropriate set-up of the IVFSs partitions by means of a post-processing tuning step applied to FRBCSs. On the other hand, we have proposed a methodology that simultaneously tunes both the degree of ignorance and the lateral position of each linguistic label, which leads to a better contextualization of the data base.

From the experimental study carried out, we must highlight the good performance of our cooperative evolutionary tuning model since it enhances the performance of the approaches with and without IVFSs, that is: the basic FH-GBML FRBCS, the initial FRBCS post-processed with lateral tuning and the two tuning approaches that compose our model when they are applied separately. The results obtained with this methodology allow us to determine, with the corresponding statistical support, that the combination of the selected tuning approaches is very useful to solve classification tasks by tuning the semantic of the fuzzy partitions in a more optimal way. Furthermore, the enhancement of the results provided by the cooperative evolutionary tuning model, especially with respect to the results of the tuning model based on the linguistic 3-tuples representation, reflects a good management of the uncertainties of the system leading to a global performance improvement.

Acknowledgements

This work was supported in part by the Spanish Ministry of Science and Technology under projects TIN2008-06681-C06-01 and TIN2010-15055.

References

[1] M. R. Akbarzadeh-Totonchi and M. Moshtagh-Khorasani, "A hierarchical fuzzy rule-based approach to aphasia diagnosis", *J. Biomed. Inform.* **40**(5) (2007) 465–475.

[2] R. Alcalá, J. M. Benítez, J. Casillas, O. Cordón and R. Pérez, "Fuzzy control of HVAC systems optimized by genetic algorithms", *Applied Intelligence* **18** (2003) 155 – 177

[3] R. Alcalá, J. Alcalá-Fdez, M. Gacto and F. Herrera, "Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation", *Soft Comput.* **11**(5) (2007) 401 – 419.

[4] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto and F. Herrera, "Improving fuzzy logic controllers obtained by experts: a case study in HVAC systems", *Applied Intelligence* **31**(1) (2007) 15 – 30

[5] R. Alcalá, J. Alcalá-Fdez and F. Herrera, "A proposal for the Genetic Lateral Tuning of Linguistic Fuzzy Systems and its Interaction with Rule Selection", *IEEE T Fuzzy Syst.* **15**(4) (2007) 616–635

[6] J. Alcalá-Fdez, L. Sánchez, S. García, M. J. del Jesus, S. Ventura, J. M. Garrell, J. Otero, J. Bacardit, V. M. Rivas, J. C. Fernández and F. Herrera, "KEEL: a software tool to assess evolutionary algorithms for data mining problems", *Soft Comput.* **13**(3) (2009) 307–318

[7] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez and F. Herrera, "KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework", *J. Mult-Valued Log. S.* in Press, Corrected Proof, 2010.

[8] M. Antonelli, P. Ducange, B. Lazzerini and F.Marcelloni, "Learning concurrently partition granularities and rule bases of Mamdani fuzzy systems in a multi-objective evolutionary framework", *Int. J Approx. Reason.* **50**(7) (2009) 1066–1080

[9] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets", *Int. J Approx. Reason.* **23** (2000) 137–209.

[10] H. Bustince, E. Barrenechea, M. Pagola and J. Fernandez, "Interval-valued fuzzy sets constructed from matrices: Application to edge detection", *Fuzzy Set Syst.* **160**(13) (2009) 1819–1840.

[11] H. Bustince, M. Pagola, E. Barrenechea, J. Fernandez, P. Melo-Pinto, P. Couto, H.R. Tizhoosh and J. Montero, "Ignorance functions. An application to the calculation of the threshold in prostate ultrasound images", *Fuzzy Set Syst.* **161**(1) (2010) 20 - 36.

[12] J. Casillas, O. Cordón, M. J. del Jesus and F. Herrera, "Genetic tuning of fuzzy rule deep structures preserving interpretability for linguistic modeling", *IEEE T Fuzzy Syst.* **13**(1) (2005) 13–29.

[13] N. R. Cazarez-Castro and L. T. Aguilar and O. Castillo, "Fuzzy logic control with genetic membership function parameters optimization for the output regulation of a servomechanism with nonlinear backlash", *Expert Syst. Appl.* **37**(6) (2010) 4368 – 4378.

[14] A. Celikyilmaz and I. B. Turksen, "Decision making with imprecise parameters", *Int. J Approx. Reason.* **51**(8) (2010) 869 – 882.

[15] S. Coupland and R. John, "New geometric inference techniques for type-2 fuzzy sets", *Int. J Approx. Reason.* **49**(1) (2008), 198 – 211.

[16] O. Cordón, M. J. del Jesus and F. Herrera, "A proposal on reasoning methods in fuzzy rule-based classification systems", *Int. J Approx. Reason.* **20**(1) (1999), 21–45.

[17] O. Cordón, F. Herrera, F.Hoffmann and L. Magdalena, *Genetic Fuzzy Systems. Evolutionary Tuning and Learning of Fuzzy Knowledge Bases.* (World Scientific, Singapore, 2001)

[18] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann and L. Magdalena, "Ten years of genetic fuzzy systems: Current framework and new trends", *Fuzzy Set Syst.* **141**(1) (2004) 5–31.

[19] C. Cortes and V. Vapnik, "Support vector networks", *Mach. Learn.* **20** (1995) 273–297

[20] L. delaOssa, J. A. Gámez and J. M. Puerta, "Learning weighted linguistic fuzzy rules by using specifically-tailored hybrid estimation of distribution algorithms", *Int. J Approx. Reason.* **50**(3) (2009) 541 - 560

[21] J. Demšar, "Statistical comparisons of classifiers over multiple data sets", *J Mach. Learn. Res.*, **7** (2006) 1–30.

[22] G. Deschrijver, "Generalized arithmetic operators and their relationship to t-norms in interval-valued fuzzy set theory", *Fuzzy Set Syst.* **160**(21) (2009) 3080–3102.

[23] L. J. Eshelman, *The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination*, (Foundations of Genetic Algorithms. Morgan Kaufman, Ch., 1991) pp. 265–283.

[24] A. Fernández, M. J. del Jesus and F. Herrera, "Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets", *Int. J Approx. Reason.* **50**(3) (2009) 561 - 577

[25] A. Fernández, M. J. del Jesus and F. Herrera, "On the 2-tuples based genetic tuning performance for fuzzy rule based classification systems in imbalanced data-sets", *Inform. Sciences* **180**(8) (2010) 1268–1291

[26] M. J. Gacto, R. Alcalá and F. Herrera, "Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems", *Soft Comput.* **13**(3) (2009) 419 – 436

[27] S. García and F. Herrera, "An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons", *J Mach. Learn. Res.*, **9** (2008) 2677–2694.

[28] S. García, A. Fernández, J. Luengo and F. Herrera, "A study of statistical techniques and performance measures for genetics–based machine learning: Accuracy and interpretability", *Soft Comput.* **13**(10) (2009) 959–977.

[29] S. García, A. Fernández, J. Luengo and F. Herrera, "Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power", *Inform. Sciences* **180**(10) (2010) 2044–2064.

[30] M. B. Gorlzakczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets", *Fuzzy Set Syst.* **21** (1987) 1–17.

[31] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words", *IEEE T Fuzzy Syst.* **8**(6) (2000) 746–752.

[32] F. Herrera, M. Lozano and A. M. Sánchez, "A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study", *Int. J Intell. Syst.* **18** (2003) 309–338.

[33] F. Herrera, "Genetic fuzzy systems: Taxonomy, current research trends and prospects", *Evolutionary Intelligence* **1** (2008) 27–46.

[34] D. Hidalgo, O. Castillo and P. Melin, "Type-1 and type-2 fuzzy inference systems as integration methods in modular neural networks for multimodal biometry and its optimization with genetic algorithms", *Inform. Sciences* **179**(13) (2009) 2123 – 2145.

[35] H. Ishibuchi and T. Nakashima, "Effect of rule weights in fuzzy rule-based classification systems", *IEEE T Fuzzy Syst.* **9**(4) (2001) 506–515.

[36] H. Ishibuchi, T. Nakashima and M. Nii, *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining.* (Springer-Verlag, Berlin, 2004)

[37] H. Ishibuchi, T. Yamamoto and T. Nakashima, "Hybridization of fuzzy GBML approaches for pattern classification problems", *IEEE T Syst. Man Cy B* **35**(2) (2005) 359–365.

[38] H. Ishibuchi and T. Yamamoto, “Rule weight specification in fuzzy rule-based classification systems”, *IEEE T Fuzzy Syst.* **13** (2005) 428–435

[39] M. Kaya and R. Alhajj, “Utilizing genetic algorithms to optimize membership functions for fuzzy weighted association rules mining”, *Applied Intelligence* **24**(1) (2006) 7 – 15

[40] Q. Liang and J. M. Mendel, “Interval type-2 fuzzy logic systems: Theory and design”, *IEEE T Fuzzy Syst.* **8** (2000) 535–549.

[41] J. Luengo and F. Herrera, “Domains of competence of fuzzy rule based classification systems with data complexity measures: A case of study using a fuzzy hybrid genetic based machine learning method”, *Fuzzy Set Syst.* **161**(1) (2010) 3–19

[42] E.G. Mansoori, M.J. Zolghadri and S.D. Katebi, “Protein Superfamily Classification Using Fuzzy Rule-Based Classifier”, *IEEE T Nanobiosci* **8**(1) (2009) 92–99

[43] S. Miller and R. John, “An Interval Type-2 Fuzzy multiple echelon supply chain model”, *Knowl-Based Syst.* **23**(4) (2010) 363–368

[44] Y. Nojima, H. Ishibuchi and I. Kuwajima, “Parallel distributed genetic fuzzy rule selection”, *Soft Comput.* **13**(3) (2009) 511 – 519

[45] A. M. Palacios, L. Sánchez and I. Couso, “Diagnosis of dyslexia with low quality data with genetic fuzzy systems”, *Int. J Approx. Reason.* **51**(8) (2010) 993–1009

[46] B.-J. Park, W. Pedrycz and S.-K Oh, “Simplified fuzzy inference rule-based genetically optimized Hybrid Fuzzy Neural Networks”, *Int. J. Uncertain Fuzz.* **16**(2) (2008) 245–274

[47] J. Platt, “Fast Training of Support Vector Machines Using Sequential Minimal Optimization”, in *Advances in Kernel Methods – Support Vector Learning*. B. Schlkopf and C. Burges and A. Smola, eds. (MIT Press, Cambridge MA, 1998) 42–65

[48] R. Sambuc, “Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne”, Ph.D. thesis, University of Marseille, 1975.

[49] J. Sanz, A. Fernandez, H. Bustince and F. Herrera, “Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning”, *Inform. Sciences*, **180**(19) (2010) 3674–3685

[50] G. Schaefer, M. Závisek and T. Nakashima, “Thermography based breast cancer analysis using statistical features and fuzzy classification”, *Pattern Recogn.* **42**(6) (2009) 1133 – 1137

[51] D. Sheskin, *Handbook of parametric and nonparametric statistical procedures*, (Chapman & Hall/CRC, second edition, 2006)

[52] J. T. Starczewski, "Efficient triangular type-2 fuzzy logic systems", *Int. J Approx. Reason.* **50**(5) (2009), 799 – 811.

[53] V. Vapnik, *Statistical Learning Theory.* (Wiley, New York, U.S.A., 1998)

[54] C. L. Walker and E. A. Walker, "Sets with type-2 operations", *Int. J Approx. Reason.* **50**(1) (2009) 63 – 71.

[55] D. Wu and W. W. Tan, "Genetic learning and performance evaluation of interval type-2 fuzzy logic controllers", *Eng. Appl. Artif. Intel.* **19**(8) (2006) 829 – 841.

[56] D. Wu and J. M. Mendel, "Uncertainty measures for interval type-2 fuzzy sets", *Inform. Sciences* **177**(23) (2007) 5378 – 5393.

[57] M. F. Zarandi, A. Alaeddini and I. Turksen, "A hybrid fuzzy adaptive sampling - run rules for shewhart control charts", *Inform. Sciences* **178**(4) (2008) 1152–1170.